

VU Research Portal

Empirical Analysis of Affine vs. Nonaffine Variance Specifications in Jump-Diffusion Models for Equity Indices

Seeger, N.J.; Rodrigues, P.J.M.; Ignatieva, K.

published in

Journal of Business and Economic Statistics
2015

DOI (link to publisher)

[10.1080/07350015.2014.922471](https://doi.org/10.1080/07350015.2014.922471)

document version

Early version, also known as pre-print

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Seeger, N. J., Rodrigues, P. J. M., & Ignatieva, K. (2015). Empirical Analysis of Affine vs. Nonaffine Variance Specifications in Jump-Diffusion Models for Equity Indices. *Journal of Business and Economic Statistics*, 33(1), 68-75. <https://doi.org/10.1080/07350015.2014.922471>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Empirical Analysis of Affine vs. Nonaffine Variance Specifications in Jump-Diffusion Models for Equity Indices

Katja IGNATIEVA

School of Risk and Actuarial Studies, Australian School of Business,
University of New South Wales, Sydney, Australia, NSW 2052
(k.ignatieva@unsw.edu.au)

Paulo RODRIGUES

Finance Department, Maastricht University, Maastricht, Netherlands, 6211 LM,
(p.rodriques@maastrichtuniversity.nl)

Norman SEEGER

Finance Department, VU University Amsterdam, Amsterdam, Netherlands, 1084 HV,
(n.j.seeger@vu.nl)

This version: May 4, 2014

Abstract This paper investigates several crucial issues that arise when modeling equity returns with stochastic variance. (i) Does the model need to include jumps even when using a nonaffine variance specification? We find that jump models clearly outperform pure stochastic volatility models. (ii) How do affine variance specifications perform when compared to nonaffine models in a jump diffusion setup? We find that nonaffine specifications outperform affine models, even after including jumps.

Key Words Bayesian inference; Deviance information criteria; Jump diffusion; Markov Chain Monte Carlo; Stochastic volatility.

JEL Classifications: C11, G11, G12

Empirical Analysis of Affine vs. Nonaffine Variance Specifications in Jump-Diffusion Models for Equity Indices

This version: May 4, 2014

Abstract This paper investigates several crucial issues that arise when modeling equity returns with stochastic variance. (i) Does the model need to include jumps even when using a nonaffine variance specification? We find that jump models clearly outperform pure stochastic volatility models. (ii) How do affine variance specifications perform when compared to nonaffine models in a jump diffusion setup? We find that nonaffine specifications outperform affine models, even after including jumps.

Key Words Bayesian inference; Deviance information criteria; Jump diffusion; Markov Chain Monte Carlo; Stochastic volatility.

JEL Classifications: C11, G11, G12

1. INTRODUCTION

The objective of this paper is to investigate and compare two strands of the literature concerned with modeling equity returns. On the one hand, [Jones \(2003b\)](#), [Bakshi, Ju, and Ou-Yang \(2006\)](#), [Christoffersen, Jacobs, and Mimouni \(2010\)](#), and [Chourdakis and Dotsis \(2011\)](#) propose drift and diffusion specifications for the variance process that result in a nonaffine model framework. However, in these papers, neither jumps in returns nor jumps in variance are considered important model components. In contrast, [Eraker, Johannes, and Polson \(2003\)](#) and [Broadie, Chernov, and Johannes \(2007\)](#) include jump components in the return and variance process, but preserve the drift and the diffusion term of the variance process in such a way that their models continue to be affine models.

The difference between these specifications is important. Both approaches employ a stochastic volatility term, but use different procedures to model sudden large movements in the price and variance process. The nonaffine specification introduces nonlinearities into the drift and diffusion term of the variance process. This provides more flexibility when attempting to capture large sudden movements in returns and variance, compared with the linear structure of an affine specification. The second approach generates large movements in returns and variance by staying within an affine framework and assuming a discontinuity in the price and, possibly, the variance process modeled via a jump component. The advantage of staying within the affine model class, analyzed in [Duffie, Pan, and Singleton \(2000\)](#), is that it allows for quasi-closed-form solutions for European option prices, portfolio rules, and the computation of transition densities, all of which are advantageous for practical implementation. Furthermore, the mathematical properties of these models are well understood. The nonaffine models lack these benefits.

Our empirical analysis is designed to answer the question of whether we can disregard a jump component when modeling the variance process in a nonaffine way? Empirical studies, e.g. by [Jones \(2003b\)](#), [Bakshi et al. \(2006\)](#), and [Chourdakis and Dotsis \(2011\)](#) propose nonaffine models that ignore jumps, and model extreme return and variance movements by means of nonlinear structures

in the drift and diffusion component of the variance process. However, there is a strand of literature in which a nonparametric model setup is used to analyze whether jumps in returns and jumps in variance are important model components; see, e.g. [Lee and Mykland \(2008\)](#), [Aït-Sahalia and Jacod \(2010\)](#), [Barndorff-Nielsen and Shephard \(2006\)](#), [Corsi, Pirino, and Renò \(2010\)](#), and [Dumitru and Urga \(2012\)](#). These studies find that jumps are indeed an important model component. Motivated by these results, we analyze, in a parametric model setup, whether nonlinearities and jumps are in fact substitutes for each other or whether including jumps can improve model performance even beyond the introduction of nonlinearities. Our findings clearly show that jumps do indeed play an important role in modeling the return process of a major equity index like the S&P 500 even after leaving the affine model class. Specifically, we find that even the worst performing jump model outperforms every pure stochastic volatility model considered.

A second, subordinate question analyzed is whether a nonaffine specification outperforms an affine specification within a jump diffusion setting. It has been shown in the literature, e.g. by [Jones \(2003b\)](#), [Bakshi et al. \(2006\)](#), [Christoffersen et al. \(2010\)](#), and [Chourdakis and Dotsis \(2011\)](#) that this is indeed the case in a pure stochastic volatility setting. This question relates to the principle of keeping a model specification as simple as possible. The higher computational costs of using a nonaffine model instead of the more familiar affine specification can be justified only if the nonaffine specification results in significantly better model performance. With regard to this question, we find that nonaffine specifications do perform considerably better than affine setups even after including a jump term.

2. MODEL DESCRIPTION

We assume that the logarithm of the index price $Y_t = \ln(S_t)$ and the variance V_t solve the following system of stochastic differential equations:

$$dY_t = \mu dt + \sqrt{V_t} dW_t^y + d\left(\sum_{j=1}^{N_t} \xi_j^y\right) \quad (1)$$

$$dV_t = \left(\alpha_0 + \alpha_1 \frac{1}{V_t} + \alpha_2 V_t + \alpha_3 V_t^2\right) dt + \sigma_v V_t^b dW_t^v + d\left(\sum_{j=1}^{N_t} \xi_j^v\right) \quad (2)$$

where dW_t^y and dW_t^v denote Brownian increments with correlation $E(dW_t^y dW_t^v) = \rho dt$, with ρ modeling the so-called leverage effect. The term μ captures the expected return, N_t denotes a Poisson process with constant intensity λ . The Poisson process enters both the return and variance equation, thus generating simultaneous jumps. The parameters ξ^y and ξ^v denote respectively jump sizes in returns and variance. We assume that the jump sizes in variance and returns are correlated. The jump size in variance follow an exponential distribution with expectation μ_v , i.e., $\xi_t^v \sim \text{Exp}(\mu_v)$, and the jump sizes in returns follow a conditional normal distribution with mean given by $\mu_y + \rho_j \xi_t^v$ and variance given by σ_y^2 , i.e., $\xi_t^y | \xi_t^v \sim N(\mu_y + \rho_j \xi_t^v, \sigma_y^2)$.

This model setup results in a general and flexible framework that subsumes a large number of modeling approaches proposed and tested independently in the related literature. By restricting the full model setup to various special cases, we can compare the performance of the different approaches to modeling the evolution of equity dynamics.

The return process in Equation (1) is based on [Bates \(1996\)](#), in which a combination of a stochastic volatility and jumps in returns model is used to analyze exchange rate processes. That is, for modeling jumps we employ a jump model with stochastic jump sizes and constant jump intensity, a framework that is frequently used in the literature; see e.g. [Eraker et al. \(2003\)](#), [Broadie et al. \(2007\)](#), [Durham \(2013\)](#), or [Ferriani and Pastorello \(2012\)](#). The variance process in Equation (2) nests

several specifications used in the literature. In the full model the drift component of the variance process follows a polynomial specification (POLY) that is applied in, e.g. [Conley, Hansen, Luttmer, and Scheinkman \(1997\)](#) and [Aït-Sahalia \(1996\)](#) to analyze short rate models and in [Chourdakis and Dotsis \(2011\)](#) to analyze stochastic volatility models without a jump component in the context of equity returns. The diffusion part of the variance is modeled as a constant elasticity of variance (CEV) specification where the exponent parameter b of the variance is undetermined and is estimated freely. [Jones \(2003b\)](#) uses this specification to analyze equity indices. Modeling jumps in variance follows the specification given in [Eraker et al. \(2003\)](#).

To answer our research questions, we compare the full unrestricted model described in Equation (1) and Equation (2) with several restricted specifications. We differentiate three main model classes in terms of how jump components are treated. Models with jumps in returns and variance are assigned to the stochastic volatility with correlated jumps (SVCJ) model class. Models that keep jumps in returns but remove them from the variance process are in the SVJ model class. Finally, models that switch off jumps in both the return and the variance process are in the pure stochastic volatility (SV) model class. Within each model class, we further classify models according to how the drift and diffusion components of the variance process are specified. Restricting the POLY drift component in Equation (2) by setting α_1 and α_3 equal to zero, we obtain an affine linear drift specification (ALIN). For the diffusion component in Equation (2), [Christoffersen et al. \(2010\)](#) suggest specifications in which the CEV exponent parameter b is restricted to 0.5, 1, or 1.5, termed SQR, ONE, or 3/2, respectively. Combinations of the aforementioned restrictions make up many well-known model specifications. For example, not considering jumps and combining an ALIN drift specification with a square root diffusion specification, SQR, results in the famous SV model of [Heston \(1993\)](#).

Having three model classes (SV, SVJ, SVCJ), two variance drift specifications (POLY, ALIN), and four variance diffusion specifications (CEV, SQR, ONE, 3/2) results in a total of 24 models to analyze, which are listed in Table 1. However, in order to avoid overloading the paper with too many results, we only present results in detail that are useful for answering our research questions. The

Table 1. Overview of Models

Drift	Diffusion	Features
Affine	SQR	variance drift is affine in variance, square root diffusion
Affine	VAR	variance drift is affine in variance, linear diffusion
Affine	3/2	variance drift is affine in variance, 3/2 diffusion
Affine	CEV	variance drift is affine in variance, free diffusion
POLY	SQR	variance drift is polynomial in variance, square root diffusion
POLY	VAR	variance drift is polynomial in variance, linear diffusion
POLY	3/2	variance drift is polynomial in variance, 3/2 diffusion
POLY	CEV	variance drift is polynomial in variance, free diffusion

NOTE: This table shows the different specifications of the drift and diffusion terms for the dynamics of the stochastic variance. For each model class SV, SVJ, and SVCJ, we estimate every specification given in the table.

decision on which results to present is based on model performance, and statistical and economic reasoning. The complete results are available from the authors on request.

3. EMPIRICAL IMPLEMENTATION

3.1 Discretization

To estimate the model, we use an Euler discretization scheme and set the interval at $\Delta = 1$, which corresponds to one day. Denoting $R_t = Y_t - Y_{t-1}$ as the log-return of the asset, we can write the discretized version of the system of equations given in (1) and (2) as

$$\begin{aligned}
 R_t &= \mu + \sqrt{V_{t-1}}\varepsilon_t^y + \xi_t^y J_t \\
 V_t &= V_{t-1} + \alpha_0 + \alpha_1 \frac{1}{V_{t-1}} + \alpha_2 V_{t-1} + \alpha_3 V_{t-1}^2 + \sigma_v V_{t-1}^b \varepsilon_t^v + \xi_t^v J_t,
 \end{aligned} \tag{3}$$

where shocks to returns and volatility, $\varepsilon_t^y = W_t^y - W_{t-1}^y$ and $\varepsilon_t^v = W_t^v - W_{t-1}^v$, follow a bivariate normal distribution with zero expectation, unit variance, and correlation ρ . In the Euler discretization

scheme, we assume at most one jump per day because, given the observation frequency, we cannot distinguish between “more jumps” and “bigger jumps”. That is, the indicator J_t is set equal to one in the event a jump occurs and equal to zero in the case of no jump. Note that the jump indicator J_t in the return equation is identical to the indicator in the variance equation since jumps occur simultaneously. The jump sizes retain the distributional assumptions described in Section 2.

The assumption of at most one jump per day could lead to some discretization bias when estimating jump parameters. However, the following example demonstrates that since jumps are rare events, discretization bias is typically very small. Using $P(N_t - N_{t-1} = j) = \frac{\exp\{-\lambda\}\lambda^j}{j!}$ and assuming the jump intensity to be $\lambda = 0.1$, the probability of observing more than one jump per day is 0.0047. Note that our estimation results indicate estimates for λ much smaller than 0.1.

For technical details concerning the discretization schemes and the existence of stationary distributions of the models, as well as simulation results, the reader is referred to [Jones \(2003b\)](#), [Eraker et al. \(2003\)](#), [Jones \(2003a\)](#), [Aït-Sahalia \(1996\)](#), and [Conley et al. \(1997\)](#).

3.2 Estimation

We employ a Bayesian estimation and model testing strategy for our empirical investigation. These procedures were first used in the context of estimating continuous time models for equity returns by [Jacquier, Polson, and Rossi \(1994\)](#) and [Jacquier, Polson, and Rossi \(2004\)](#). Since then, these methods have been successfully employed in many empirical studies. In the following section, we provide an overview of the sampling algorithm for a SVCJ model, since this is the most complex setup used in our analysis. Estimation of the restricted models accordingly follows. For a general introduction to MCMC methods, the reader is referred to [Casella and George \(1992\)](#), [Chib and Greenberg \(1995\)](#), and [Johannes and Polson \(2009\)](#).

Bayes Theorem implies that the posterior distribution of the parameters and the latent states is proportional to the likelihood times the prior distribution. Using the Euler discretized version of the

models, it follows that the posterior distribution is proportional to

$$\prod_{t=1}^T p(R_t, V_t | V_{t-1}, \xi_t^y, \xi_t^v, \Theta_{RV}) p(\xi_t^y | \xi_t^v, \mu_y, \rho_J, \sigma_y) p(\xi_t^v | \mu_v) p(J_t | \lambda) p(\Theta) \quad (4)$$

where $\Theta_{RV} = \{\mu, \rho, \alpha_0, \alpha_1, \alpha_2, \alpha_3, \sigma_v, b\}$ denotes the parameter vector of the joint distribution of returns and variance and Θ denotes the full parameter vector comprising all parameters of the model. We assume independent conjugate priors for the parameters, so that the prior for the full parameter vector $p(\Theta)$ can be decomposed into the product of these priors. The prior parameters are set in accordance with [Eraker et al. \(2003\)](#), [Jacquier et al. \(2004\)](#), and [Li, Wells, and Yu \(2008\)](#). In our application, we use a sampler that draws one parameter and latent state at a time. Therefore, it is convenient to decompose the bivariate distribution of return and variance in Equation (4) into the product of a conditional and a marginal distribution. Given the Euler discretization, the bivariate distribution is normal with mean vector \mathbf{m} and covariance matrix \mathbf{s} given by

$$\mathbf{m} = \begin{pmatrix} \mu + \xi_t^y J_t \\ V_{t-1} + a(V_{t-1}) + \xi_t^v J_t \end{pmatrix} \quad \mathbf{s} = \begin{pmatrix} V_{t-1} & \rho \sigma_v V_{t-1}^{b+0.5} \\ \rho \sigma_v V_{t-1}^{b+0.5} & \sigma_v^2 V_{t-1}^{2b} \end{pmatrix}$$

with $a(V_{t-1}) = \alpha_0 + \alpha_1 \frac{1}{V_{t-1}} + \alpha_2 V_{t-1} + \alpha_3 V_{t-1}^2$. Given \mathbf{m} and \mathbf{s} , further decomposition of the bivariate distribution into the product of a conditional distribution times a marginal distribution is straightforward.

To derive the complete conditionals for a parameter, we simply remove all multiplicative factors from the posterior given in (4) that do not contain the parameter itself. Following this approach, the complete conditional for the parameter μ is proportional to

$$\prod_{t=1}^T p(R_t | V_t, V_{t-1}, \xi_t^y, \xi_t^v, \Theta_{RV}) p(\mu),$$

where the conditional distribution is normal with parameters

$$\begin{aligned}\mu_{R|V} &= \mu + \xi_t^y J_t + \frac{\rho V_{t-1}^{0.5-b}}{\sigma_v} (V_t - V_{t-1} - a(V_{t-1}) - \xi_t^v J_t) \\ \sigma_{R|V}^2 &= V_{t-1}(1 - \rho^2),\end{aligned}$$

and $p(\mu)$ is a conjugate prior that we assume to be $N(0, 1)$. Using $R_t = \mu_{R|V} + \sigma_{R|V} \cdot u_t$, with u_t being standard normally distributed, and rearranging terms, we derive the following regression model

$$\frac{R_t - \xi_t^y J_t - \frac{\rho V_{t-1}^{0.5-b}}{\sigma_v} (V_t - V_{t-1} - a(V_{t-1}) - \xi_t^v J_t)}{\sqrt{V_{t-1}(1 - \rho^2)}} = \frac{1}{\sqrt{V_{t-1}(1 - \rho^2)}} \mu + u_t \quad (5)$$

where μ can be interpreted as a regression coefficient. As such, we use its full conditional distribution in an MCMC step to draw μ .

The complete conditional for α_0 is proportional to

$$\prod_{t=1}^T p(V_t | R_t, V_{t-1}, \xi_t^y, \xi_t^v, \Theta_{RV}) p(\alpha_0)$$

where the conditional distribution is normal with parameters given by

$$\begin{aligned}\mu_{V|R} &= V_{t-1} + a(V_{t-1}) + \xi_t^v J_t + \rho \sigma_v V_{t-1}^{b-0.5} (R_t - \mu - \xi_t^y J_t) \\ \sigma_{V|R}^2 &= \sigma_v^2 (1 - \rho^2) V_{t-1}^{2b}\end{aligned}$$

and $p(\alpha_0)$ is a conjugate prior that we assume to be $N(0, 1)$. Using the conditional normal property and by rearranging terms we can derive the following regression model

$$\frac{V_t - V_{t-1} - b(V_{t-1}) + \xi_t^v J_t - \rho \sigma_v V_{t-1}^{b-0.5} (R_t - \mu - \xi_t^y J_t)}{\sigma_v \sqrt{(1 - \rho^2) V_{t-1}^b}} = \frac{1}{\sigma_v \sqrt{(1 - \rho^2) V_{t-1}^b}} \alpha_0 + u_t \quad (6)$$

where $b(V_{t-1}) = \alpha_1 \frac{1}{V_{t-1}} - \alpha_2 V_{t-1} - \alpha_3 V_{t-1}^2$ and $u_t \sim N(0, 1)$. Again, by using the regression

model in Equation (6) and a standard normally distributed prior, we can derive the parameter of the complete conditional distribution for α_0 . Estimation of the remaining α parameters consists simply of varying the regression setup of equation (6), where we assume identical prior distributions for all α parameters.

To draw the parameters ρ and σ_v , we follow [Jacquier et al. \(2004\)](#) and define $\phi \equiv \rho\sigma_v$ and $\omega^2 \equiv \sigma_v^2(1 - \rho^2)$. Drawing ϕ and ω^2 amounts to using a regression setup given by

$$\frac{V_t - V_{t-1} - a(V_{t-1}) - \xi_t^v J_t}{V_{t-1}^b} = V_{t-1}^{-0.5} (R_t - \mu - \xi_t^y J_t) \phi + \omega u_t$$

where we use $\phi|\omega^2 \sim N(0, 1/2\omega^2)$ and $\omega^2 \sim IG(2, 200)$ as prior distributions, and IG denotes an inverse gamma distribution.

The complete conditional distribution for the jump parameters μ_y , ρ_J , and σ_y is given by

$$\prod_{t=1}^T p(\xi_t^y | \xi_t^v, \mu_y, \rho_J, \sigma_y) p(\mu_y) p(\rho_J) p(\sigma_y)$$

where the priors assumed are given by $\mu_y \sim N(0, 100)$, $\rho_J \sim N(0, 4)$, and $\sigma_y \sim IG(5, 1/20)$. To draw the three parameters, the following regression setup is used

$$\xi_t^y = \mu_y + \rho_J \xi_t^v + \sigma_y \varepsilon.$$

For the parameter μ_v , the complete conditional takes the form

$$\prod_{t=1}^T p(\xi_t^v | \mu_v) p(\mu_v)$$

where we take $IG(10, 1/10)$ as the prior distribution for μ_v . Standard results given in e.g. [Bernardo and Smith \(1995\)](#) show that the resulting complete conditional for μ_v is Inverse Gamma.

For the parameter λ , we assume a Beta distribution ($B(2, 40)$) as prior. The complete conditional

given by

$$\prod_{t=1}^T p(J_t|\lambda)p(\lambda) \quad (7)$$

is a combination of a binomial distribution and a Beta distribution. Standard results given in e.g. [Bernardo and Smith \(1995\)](#) show that the complete conditional follows a Beta distribution.

For the parameter b , we discretize the space into bins with equal prior probability, i.e. we assume $b = \{0.5, 0.502, 0.504, \dots, 2.5\}$. The complete conditional for b follows a multinomial distribution where the probability for each bin is proportional to

$$\prod_{t=1}^T p(V_t|R_t, V_{t-1}, \xi_t^y, \xi_t^v, \Theta_{RV}^i)p(b^i) \quad (8)$$

where the superscript i indicates that expression (8) is evaluated at the value that b has for the respective bin i . We note that in this particular case, the prior probabilities will drop out of the calculation of the complete conditional, since they are identical for all bins.

The state variables, J_t, ξ_t^y, ξ_t^v and V_t are drawn sequentially for each t . The jump indicator J_t follows a binomial distribution with complete conditional probabilities proportional to $\lambda p(R_t, V_t|V_{t-1}, \xi_t^v, \xi_t^y, \Theta)$ for $J_t = 1$ and $(1 - \lambda)p(R_t, V_t|V_{t-1}, \xi_t^v, \xi_t^y, \Theta)$ for $J_t = 0$.

For the jump sizes in returns, the complete conditional for ξ_t^y is proportional to

$$p(R_t|V_t, V_{t-1}, \xi_t^y, \xi_t^v, \Theta_{RV})p(\xi_t^y|\xi_t^v, \mu_y, \rho_J, \sigma_y)$$

which is normally distributed with parameters that are easy, albeit tedious, to compute. For the case of $J_t = 0$, we simply draw from the prior distribution, since the data provide no information.

The derivation of the complete conditional distribution for the jump sizes in variance follows similar

lines. The complete conditional is given by

$$p(V_t|R_t, V_{t-1}, \xi_t^y, \xi_t^v, \Theta_{RV})p(\xi_t^v|\mu_v)$$

which results in a truncated normal distribution with parameters that are straightforward to compute.

In the case of $J_t = 0$, we draw from the prior, since the data provide no information.

For V_t , the complete conditional is given by

$$p(R_{t+1}, V_{t+1}|V_t, J_{t-1}, \xi_{t+1}^v, \xi_{t+1}^y, \Theta_{RV})p(V_t|R_t, V_{t-1}, \xi_t^v, J_t, \Theta_{R,V})$$

which does not resemble any known statistical distribution. We therefore use the Metropolis-Hastings step to draw variances from their complete conditional distributions. We follow [Chib and Greenberg \(1995\)](#) and [Jones \(2003b\)](#) in using the recognizable part of the complete conditional as the proposal density, i.e., we draw V_t from $p(V_t|R_t, V_{t-1}, \xi_t^v, J_t, \Theta_{R,V})$. The acceptance probability for the draw is given by

$$\min \left\{ \frac{p(R_{t+1}, V_{t+1}|V_t^{(g)}, J_{t-1}, \xi_{t+1}^v, \xi_{t+1}^y, \Theta_{RV})}{p(R_{t+1}, V_{t+1}|V_t^{(g-1)}, J_{t-1}, \xi_{t+1}^v, \xi_{t+1}^y, \Theta_{RV})}, 1 \right\}$$

where $V_t^{(g)}$ indicates the proposed value and $V_t^{(g-1)}$ denotes the current draw. That is, by using V_{t-1} and R_t to obtain candidate draws, we utilize information in the data to increase the acceptance probability. Our acceptance rates are around 65%, which is nearly identical to the rates reported in [Jones \(2003b\)](#).

3.3 Model Comparison

Comparing and then ranking the models in our model setup is not easily accomplished. In theory, the most convincing statistics for model comparison are Bayes factors. But although they are theoretically appealing, these statistics are difficult to compute for very high dimensional problems like those

under consideration. In certain cases, for example comparing jump models to models without jumps, [Eraker et al. \(2003\)](#) show how to use the structure of these nested models to compute the Bayes factors. For these cases, we follow their procedure and so can use the Bayes factors for model comparison. To compare non-nested models, we rely on the deviance information criterion (DIC) derived in [Spiegelhalter, Best, Carlin, and van der Linde \(2002\)](#). This information criterion uses the same structure as any information criterion, namely, it penalizes model complexity and rewards model fit. The only difference is that the DIC accounts for the hierarchical structure of our models in determining model complexity. DIC is employed to compare stochastic variance models for equity index returns by, for example, [Berg, Meyer, and Yu \(2004\)](#).

3.4 Model Implementation

Implementation of the MCMC algorithm is carried out in C++ using random number generators of the GNU Scientific Library. Since the MCMC method is by construction a sequential algorithm (each draw depends on the preceding draw), there is limited potential to decrease computational time by parallelizing the algorithm. However, the dependence structure of the variances can be used to draw variances in blocks of two, as is done in [Jones \(2003b\)](#), which at least offers the possibility of some performance gain by parallelization. Convergence of model parameters relies heavily on the model specification to be estimated. For a SV-ALIN-SQR model, convergence can be obtained relatively quickly by drawing 50,000 times with a burn-in period of 10,000 draws, whereas for a more complex model such as SVCJ-POLY-CEV, convergence is obtained after 300,000 draws with a burn-in period of 100,000. Note that calculating stable values for the DIC statistics and, therefore, a stable ranking of the models in terms of DIC is more sensitive to the number of MCMC draws than is parameter convergence. Thus, to ensure convergence of all models in parameters and to obtain a stable ranking in terms of DIC, we base our estimation results on 2 million draws with a burn-in period of 600,000 draws. Again, the run time for a full model estimation is strongly dependent on the model specification to be estimated. We perform our calculations on a large computer cluster

equipped with Intel Xeon L5520 2.26 GHz processors. Our setup of 2 million MCMC draws results in an estimation time of about 4.5 hours for the SV-ALIN-SQR, 5.2 hours for the SVCJ-POLY-3/2 model with fixed b , and about 28 hours for the SVCJ-POLY-CEV model with estimated b parameter.

4. RESULTS

4.1 Data and Parameter Estimates

We analyze model performance using a time series of daily log returns of the S&P 500 index. Data of S&P 500 simple returns are taken from CRSP (crsp variable name “sprtrn”), which are calculated close-to-close. Simple returns are converted to log returns and the data set covers the period from January 1980 to December 2010. The mean return for the time period is about 8% p.a. with a volatility of about 18% p.a. The S&P 500 index had negative skewness of about -1.21 and a kurtosis of about 31 indicating fat tails in the period under analysis, where both numbers refer to a daily period. To check whether our results are driven by the recent financial crisis we analyze model performance for the time period 1980 – 2007 and find no difference. These additional results are available upon request.

The parameter estimators for the models within the SVCJ class can be found in Table 2. We follow the convention of reporting our estimators in daily levels with the returns given in percentages, i.e., we estimate the models and report the results with returns defined as $100 \cdot (\ln S_t - \ln S_{t-1})$. Our estimated parameters for the models with a typical square root volatility specification are in line with what can be found in the literature, see e.g. [Eraker et al. \(2003\)](#), [Eraker \(2004\)](#), and [Christoffersen et al. \(2010\)](#). We note that in order to compare the parameter estimates with other results in the literature, it is necessary to take into account the different ways to parameterize the mean reversion component in the variance process.

Given the focus of the analysis of our paper, the estimates that warrant a more detailed discussion

are jump parameters. We find that jumps are rare and have a large negative mean. For the SVCJ model we estimate a jump intensity ranging from 0.006 (SQR) to 0.013 (3/2), which translates into roughly 1.5 to 3.3 jumps per year. If we change the variance diffusion specification from SQR to ONE to CEV to 3/2 we see that the jump intensity increases for an increase in the exponent parameter b . For the mean jump sizes in returns, standard deviation of jump sizes in returns, and the mean jump size in variance, we find decreasing absolute values for increasing exponent parameter b . These values range from -2.7 (SQR) to -1.17 (3/2) for the mean return jump size, from 2.7 (SQR) to 1.7 (3/2) for the standard deviation in the return jumps, and from 1.5 (SQR) to 0.5 (3/2) for the jump size in variance. We note that the changes in parameter estimates when changing the exponent parameter b occur monotonically. Interestingly, we find almost no differences in the estimated jump parameters for the different drift specifications ALIN and POLY. These parameters are also in line with previous literature. For example, [Eraker et al. \(2003\)](#) find a jump intensity of 0.0066 for the SVCJ-ALIN-SQR specification with a mean jump size in returns of -2.64 and a mean jump size in variance of 1.48. We observe the same effects for the SVJ model class.

However, some recent papers find jump components that do not fit these results. For example, [Ferriani and Pastorello \(2012\)](#) and [Durham \(2013\)](#) estimate very frequent jumps that have a slightly positive mean. Both papers use options, meaning that their information sets are different from ours. In particular, the paper by [Durham \(2013\)](#) discusses extensively the origins of the differences between his estimation results concerning the jump component and those from the previous literature. One of the main reasons he identifies involves the prior distribution for the jump intensity used in the previous literature, which puts very low prior probability on frequent jumps, as is also true for our setup. We therefore analyze whether our prior puts a significant restriction on the posterior results. To do so, we change the prior distribution for the jump intensity from a Beta distribution to a multinomial distribution with discrete probability mass for different parameter values. We decompose the prior parameter space for the jump intensity from 0.001 to 0.5 into intervals of 0.001, where each bin has equal prior probability, thus assuming prior ignorance on the intensity

level. In unreported results that are available on request, we find that the parameter estimates are virtually the same for both prior specifications, i.e. Beta and Multinomial prior. This holds true for all jump models under consideration. Moreover, the posterior indicates zero probability for a jump intensity larger than 0.116, which shows that our choice of 0.5 as the upper limit for the discretized parameter space of λ imposes no restriction on the posterior. We therefore conclude that our posterior estimation results are driven by data rather than by restrictions on the prior distribution. Analyzing the exact reason for the different estimation results between our setup and the setup used in [Ferriani and Pastorello \(2012\)](#) and [Durham \(2013\)](#) is beyond the scope of this paper, but presents an interesting possibility for future research.

Table 2. Parameter Estimators for the SVCJ Model Class

Parameter	ALIN-SQR	ALIN-ONE	ALIN-CEV	POLY-SQR	POLY-ONE	POLY-CEV
μ	0.0370 (0.0091)	0.0407 (0.0090)	0.0441 (0.0092)	0.0372 (0.0091)	0.0401 (0.0091)	0.0438 (0.0092)
α_0	0.0167 (0.0023)	0.0097 (0.0017)	0.0031 (0.0014)	0.0045 (0.0071)	0.0091 (0.0046)	0.0008 (0.0045)
α_1	—	—	—	0.0035 (0.0018)	0.0001 (0.0008)	0.0004 (0.0008)
α_2	-0.0227 (0.0030)	-0.0123 (0.0032)	-0.0029 (0.0036)	-0.0173 (0.0059)	-0.0101 (0.0056)	0.0005 (0.0062)
α_3	—	—	—	0.0005 (0.0007)	-0.0007 (0.0009)	-0.0008 (0.0013)
σ_v	0.1182 (0.0081)	0.1458 (0.0098)	0.1318 (0.0091)	0.1375 (0.0115)	0.1466 (0.0101)	0.1329 (0.0090)
ρ	-0.6223 (0.0376)	-0.6683 (0.0396)	-0.6979 (0.0428)	-0.6451 (0.0383)	-0.6615 (0.0392)	-0.6930 (0.0451)
μ_Y	-2.3023 (0.5830)	-1.4651 (0.4775)	-1.2326 (0.3733)	-2.0718 (0.6220)	-1.5620 (0.5066)	-1.2330 (0.3931)
ρ_J	-0.0009 (0.0233)	-0.0010 (0.0437)	-0.0012 (0.0458)	-0.0012 (0.0338)	-0.0012 (0.0437)	-0.0010 (0.0469)
σ_Y	2.5079 (0.5376)	2.0115 (0.3328)	1.7632 (0.2295)	2.7039 (0.5449)	2.0329 (0.3523)	1.7782 (0.2297)
μ_V	1.5405 (0.3283)	0.6514 (0.1539)	0.5325 (0.1056)	1.1362 (0.2577)	0.6659 (0.1670)	0.5257 (0.1100)
λ	0.0058 (0.0015)	0.0097 (0.0034)	0.0126 (0.0038)	0.0063 (0.0019)	0.0088 (0.0031)	0.0124 (0.0042)
b	—	—	1.3136 (0.0503)	—	—	1.3106 (0.0549)

NOTE: This table shows the posterior means and standard deviations (in parentheses) for the ALIN and POLY drift specification in combination with a SQR, ONE, and CEV diffusion for the SVCJ model class. The details of the models are described in Section 2. The underlying data set consists of log-returns of the S&P index for the period from January 1980 until December 2010. Parameters are presented in daily percentage units.

4.2 Model Performance

A standard result from the literature analyzing affine models is that jumps are important for explaining observed return dynamics. However, the studies analyzing nonaffine specifications, such as [Jones \(2003b\)](#), [Christoffersen et al. \(2010\)](#), [Bakshi et al. \(2006\)](#), and [Chourdakis and Dotsis \(2011\)](#), treat the jump component as at most of secondary importance, if it is even analyzed at all. There are several reasons for this quasi-disregard of the jump component in this branch of the literature. First, some authors focus on finding the best variance specification in a nonaffine setup, claiming that the results could be generalized to a setup with jumps. Second, to avoid the problem of overfitting, authors understandably wish to remain within the most parsimonious model specification, and thus focus on stochastic variance. We argue that the question of the best variance specification cannot be answered without simultaneously considering jumps, since the overall variance of a model is determined by the sum of a part driven by stochastic variance and a part driven by the jump component.

We answer the question of whether jumps are important by using our results from the DIC statistics and comparing Bayes factors. Table 3 clearly shows the dominance of jump models over pure SV specifications. All SV models are at the lower end of the table and thus 100% outperformed by the jump models. In other words, no matter how a SV model is specified, it can always be improved by including a jump component. In particular, we find that the best-performing pure SV specification (SV-POLY-SQR) does not necessarily produce the best-performing overall specification when a jump component is added. This result is strong evidence against the advisability of simply first choosing an optimal SV model and then adding a jump component.

The same result is seen from the Bayes factors shown in Table 4. Following the rule of thumb given in [Kass and Raftery \(1995\)](#) and interpreting a value greater than 6 as strong evidence in favor of the model listed first under the column heading, we see that the SV model is also rejected by the Bayes factors regardless of the specification of the variance process. Regarding the question of

Table 3. Rankings of Models by DIC

Model	DIC	pD	\bar{D}
SVCJ-POLY-3/2	18183.1	3913.9	14269.2
SVCJ-ALIN-CEV	18265.2	3729.2	14535.9
SVCJ-POLY-CEV	18313.3	3675.7	14637.6
SVCJ-ALIN-3/2	18377.9	3786.8	14591.1
SVCJ-ALIN-ONE	18558.4	3373.9	15184.4
SVCJ-POLY-ONE	18642.4	3308.3	15334.1
SVCJ-POLY-SQR	18819.6	3157.5	15662.1
SVJ-ALIN-CEV	18983.0	3202.2	15780.8
SVJ-POLY-CEV	18996.5	3192.5	15804.1
SVJ-POLY-ONE	18998.9	3032.4	15966.6
SVJ-POLY-SQR	19019.2	2928.0	16091.2
SVJ-ALIN-ONE	19025.4	3007.0	16018.5
SVCJ-ALIN-SQR	19059.8	2968.9	16090.9
SVJ-POLY-3/2	19110.6	3190.3	15920.4
SVJ-ALIN-3/2	19145.0	3141.9	16003.1
SVJ-ALIN-SQR	19221.6	2777.7	16443.8
SV-POLY-SQR	19495.2	2296.2	17199.0
SV-POLY-ONE	19612.6	2403.6	17209.0
SV-ALIN-ONE	19643.1	2369.1	17274.0
SV-POLY-CEV	19670.3	2466.1	17204.2
SV-ALIN-CEV	19679.5	2457.8	17221.7
SV-ALIN-SQR	19731.5	2221.4	17510.1
SV-POLY-3/2	19822.8	2423.0	17399.8
SV-ALIN-3/2	19829.2	2402.7	17426.6

NOTE: This table shows the DIC rankings of the various models. The second and third columns show the values for model complexity and model fit, respectively. The details of the models are described in Section 2. The underlying data set consists of log-returns of the S&P index for the period from January 1980 until December 2010.

Table 4. Bayes Factors

Drift & Diffusion	SVJ vs. SV	SVCJ vs. SV	SVCJ vs. SVJ
ALIN-SQR	31.1	8.0	-23.0
ALIN-ONE	46.6	43.5	-3.0
ALIN-3/2	49.4	45.8	-3.6
ALIN-CEV	48.9	46.1	-2.8
POLY-SQR	35.2	24.5	-10.6
POLY-ONE	46.5	41.2	-5.3
POLY-3/2	49.6	47.3	-2.3
POLY-CEV	48.6	46.3	-2.4

NOTE: This table shows the Bayes factors in a comparison of nested model specifications. The first two columns show the results for comparing the SVJ and the SVCJ models respectively with the SV model specification. The third column shows the results for comparing the SVJ and the SVCJ models. A positive value for the Bayes factor shows a preference for the model mentioned first in the column heading. The underlying data set consists of log-returns of the S&P index for the period from January 1980 until December 2010.

whether jumps can be disregarded when moving away from the affine model specification, we note that the Bayes factors even increase in the nonaffine specifications.

The results of the comparison between models with jumps in returns and models containing jumps in returns and variance are mixed. We see a clear outperformance of the SVCJ model class when looking at the DIC statistics, but this is not the case when the Bayes factors are used for model comparison. When using the Bayes factors as a means of comparison, we find strong evidence in favor of a SVJ model for two cases and weak evidence in favor of SVJ models in the remaining setups.

In summary, we conclude that jumps are an important model component in explaining in-sample return dynamics regardless of whether the variance specification is affine or nonaffine.

It has been shown in the literature, see e.g. [Jones \(2003b\)](#), [Bakshi et al. \(2006\)](#), [Christoffersen et al. \(2010\)](#), and [Chourdakis and Dotsis \(2011\)](#), that nonaffine models outperform their affine counterparts when looking at a pure SV specification. An interesting question that follows from the strong outperformance of jump diffusion models when compared to pure SV models is whether

nonaffine models still outperform their affine counterparts in a jump diffusion setup. This is an important question for a number of reasons. First, affine models are relatively well known with respect to their mathematical properties. Second, the economic implications of affine models in equilibrium models are also well known. Third, the technique of [Duffie et al. \(2000\)](#) can be used to calculate semi-closed-form solutions for prices of plain-vanilla options. None of the above-mentioned points are true of nonaffine models. The question of whether the better performance of the nonaffine models compared to the affine models justifies having to deal with the greater model complexity of the former therefore becomes of interest.

According to Table 3, all top-ranked models are of the nonaffine type. The best affine specification, SVCJ-ALIN-SQR, ranks at 12th. Since it is not possible to compute standard errors or distributions of DIC, we cannot judge what constitutes a significant difference in these statistics. We therefore use an ad-hoc approach to assessing model difference by looking at percentage differences relative to the best performing model in terms of the DIC statistics. The difference between the best and the worst model is about 9%, i.e. we achieve a 9% improvement in DIC when switching from the worst to the best model. The model improvement achieved by switching from the best affine model to the best overall model is about 4.8%. These numbers show that a substantial improvement in performance can be achieved by abandoning the affine model class. When comparing affine to nonaffine specifications within each model class, we find that the affine specification ranks last for the SVJ and SVCJ model class and third to last for the SV class.

5. CONCLUSION

We conduct a comprehensive empirical analysis of continuous-time models for equity index returns with the aim of investigating the properties of several widely used model classes, specifically affine and nonaffine stochastic variance specifications augmented by jump components. The results of our analysis lead us to the following conclusions. First, jump components are an important model component regardless of whether the setup is affine or nonaffine. Even the worst performing jump

model outperforms a pure SV specification in terms of DIC statistics. Furthermore, Bayes factors strongly favor jump models over pure diffusion models regardless of the variance setup. Second, nonaffine specifications perform considerably better than affine specifications even when jump components are included into the model.

ACKNOWLEDGMENTS

We would like to thank Shakeeb Kahn (Editor), Associate Editor, and two anonymous referees, for helpful guidance and constructive comments. For insightful discussions and suggestions, we thank Michael Johannes, Christian Schlag, Raman Uppal, Grigory Vilkov, Peter Schotman, André Lucas, Joao Cocco, Joachim Grammig, Pedro Santa-Clara, and seminar participants at the 2010 Meeting of the European Finance Association, 2010 FMA European Conference, 2010 World Congress of the Bachelier Finance Society, 2009 Annual Meeting of the German Finance Association (DGF), 2009 European Meeting of the Econometric Society, 2009 Far East and South Asia Meeting of the Econometric Society, 2009 International Conference on Computing in Economics and Finance, Econometric Society Australasian Meeting in 2009, and 2009 Meeting of the Swiss Society of Economics and Statistics. Any errors are however our responsibility alone. A previous version of the paper was circulated under the title of “Stochastic Volatility and Jumps: Exponentially Affine Yes or No? An Empirical Analysis of S&P500 Dynamics.” Financial support from the German Science Foundation is gratefully acknowledged. We thank SURFsara (www.surfsara.nl) for support in using the Lisa Computer Cluster.

References

- Aït-Sahalia, Y. (1996), “Testing Continuous-time Models of the Spot Interest Rate,” *Review of Financial Studies*, 9, 385–426.
- Aït-Sahalia, Y. and Jacod, J. (2010), “Is Brownian Motion Necessary to Model High-Frequency Data?” *The Annals of Statistics*, 38, 3093–3128.
- Bakshi, G., Ju, N., and Ou-Yang, H. (2006), “Estimation of Continuous-Time Models with an Application to Equity Volatility Dynamics,” *Journal of Financial Economics*, 82, 227–249.
- Barndorff-Nielsen, O. E. and Shephard, N. (2006), “Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation,” *Journal of Financial Econometrics*, 4, 1–30.

- Bates, D. S. (1996), "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9, 69–107.
- Berg, A., Meyer, R., and Yu, J. (2004), "Deviance Information Criterion for Comparing Stochastic Volatility Models," *Journal of Business & Economic Statistics*, 22, 107–120.
- Bernardo, J. and Smith, A. (1995), *Bayesian Theory*, John Wiley and Sons.
- Broadie, M., Chernov, M., and Johannes, M. (2007), "Model Specification and Risk Premia: Evidence from Futures Options," *The Journal of Finance*, 62.
- Casella, G. and George, E. I. (1992), "Explaining the Gibbs Sampler," *The American Statistician*, 46, 167–174.
- Chib, S. and Greenberg, E. (1995), "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, 49, 327–335.
- Chourdakis, K. and Dotsis, G. (2011), "Maximum Likelihood Estimation of Non-Affine Volatility Processes," *Journal of Empirical Finance*, 18, 533–545.
- Christoffersen, P., Jacobs, K., and Mimouni, K. (2010), "Volatility Dynamics for the S&P500: Evidence from Realized Volatility, Daily Returns, and Option Prices," *Review of Financial Studies*, 23, 3141–3189.
- Conley, T. G., Hansen, L. P., Luttmer, E. G. J., and Scheinkman, J. A. (1997), "Short-Term Interest Rates as Subordinated Diffusions," *Review of Financial Studies*, 10, 525–577.
- Corsi, F., Pirino, D., and Renò, R. (2010), "Threshold Bipower Variation and the Impact of Jumps on Volatility Forecasting," *Journal of Econometrics*, 159, 276–288.
- Duffie, D., Pan, J., and Singleton, K. (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1343–1376.
- Dumitru, A.-M. and Urga, G. (2012), "Identifying Jumps in Financial Assets: A Comparison Between Nonparametric Jump Tests," *Journal of Business & Economic Statistics*, 30, 242–255.
- Durham, G. (2013), "Risk-Neutral Modelling with Affine and Non-Affine Models," *Journal of Financial Econometrics*, (forthc.).
- Eraker, B. (2004), "Do stock prices and volatility jump? Reconciling evidence from spot and option prices," *The Journal of Finance*, LIX.
- Eraker, B., Johannes, M., and Polson, N. (2003), "The Impact of Jumps in Volatility and Returns," *The Journal of Finance*, 58, 1269–1300.
- Ferriani, F. and Pastorello, S. (2012), "Estimating and Testing Non-Affine Option Pricing Models with a Large Unbalanced Panel of Options," *The Econometrics Journal*, 15, 171–203.
- Heston, S. L. (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327–343.

- Jacquier, E., Polson, N. G., and Rossi, P. E. (1994), “Bayesian Analysis of Stochastic Volatility Models,” *Journal of Business & Economic Statistics*, 12, 371–389.
- (2004), “Bayesian analysis of stochastic volatility models with fat-tails and correlated errors,” *Journal of Econometrics*, 122, 185–212.
- Johannes, M. and Polson, N. (2009), “MCMC methods for Continuous-Time Financial Econometrics,” in *Handbook of Financial Econometrics*, Vol 2, ed. Ait-Sahalia, Y. L. H., North Holland, pp. 1–72.
- Jones, C. S. (2003a), “Nonlinear Mean Reversion in the Short-Term Interest Rate,” *Review of Financial Studies*, 16, 793–843.
- (2003b), “The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets,” *Journal of Econometrics*, 116, 181–224.
- Kass, R. and Raftery, A. (1995), “Bayes Factors,” *Journal of the American Statistical Association*, 90, 773 – 795.
- Lee, S. S. and Mykland, P. A. (2008), “Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics,” *Review of Financial Studies*, 21, 2535–2563.
- Li, H., Wells, M. T., and Yu, C. L. (2008), “A Bayesian Analysis of Return Dynamics with Lévy Jumps,” *Review of Financial Studies*, 21, 2345–2378.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002), “Bayesian Measures of Model Complexity and Fit,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 583–639.